



**PRIORITY
DOCUMENT**

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

Bescheinigung

Herr Professor Dr. med. Wolf-Georg F o r s s m a n n in Hannover/Deutschland
hat eine Patentanmeldung unter der Bezeichnung 3

"Serin-Proteinase-Inhibitoren"

E50

am 23. Dezember 1997 beim Deutschen Patent- und Markenamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig die Symbole C 07 K, A 61 K und C 07 H der Internationalen Patentklassifikation erhalten.

München, den 15. Januar 1999

Deutsches Patent- und Markenamt

Der Präsident

Im Auftrag

Zeichen: 197 57 572.2

Fraust

971954de Me/Sch-gn
23. Dezember 1997

Serin-Proteinase-Inhibitoren

Die Erfindung betrifft Serin-Proteinase-Inhibitoren, cDNA kodierend für Serin-Proteinase-Inhibitoren, Arzneimittel enthaltend die Inhibitoren oder deren codierende Nucleinsäure, Verwendungen der erfindungsgemäßen Verbindungen zur Herstellung von Arzneimitteln zur Behandlung verschiedener Indikationen, Antikörper oder Antikörperfragmente gegen Epitope der erfindungsgemäßen Verbindungen, Poly- oder Oligonucleotide, die mit Genen der erfindungsgemäßen Verbindungen hybridisieren, ein Diagnos-tikum zum Aufspüren der erfindungsgemäßen Verbindungen, sowie Arzneimittel enthaltend Antikörper oder Poly- oder Oligonucleo-tide gemäß der Erfindung.

Proteolytische Prozesse spielen in allen Organismen eine bedeuti-
tende physiologische Rolle, wobei zwischen unspezifischen und
spezifischen proteolytischen Reaktionen zu unterscheiden ist.
Zu den ersten gehören beispielsweise der Nahrungsaufschluß im
Verdauungstrakt durch Endopeptidasen sowie der intrazelluläre
Abbau verbrauchter endogener Substanzen und phagozytierten
Materials durch lysosomale Proteininasen. Spezifische Proteolysen
dienen meistens der Überführung eines Proenzyms in die aktive
Form wie bei der Überführung von Trypsinogen in Trypsin und
Chymotrypsinogen in Chymotrypsin sowie bei den Kallikrein-Kinin-
Kaskaden und der Blutgerinnungskaskade. Je nach Beschaffenheit
des reaktiven Zentrums der daran beteiligten Proteininasen werden
diese in die Klassen der Serin-Proteininasen (z.B. Chymotrypsin,
Trypsin, Elastase und Kathepsin G), der Aspartat-Proteininasen
(z.B. Kathepsin D, Kathepsin E und Pepsin), der Cystein-Protei-nasen
(z.B. Kathepsin B, Kathepsin H und Kathepsin L) und der
Metallo-Proteininasen (z.B. Kollagenase und Thermolysin) unter-teilt.

Um die oft kaskadenartig verlaufenden proteolytischen Prozesse gegenregulieren zu können, verfügt der Organismus über eine Reihe von anderen Proteinen, den Proteinase-Inhibitoren (zur Übersicht siehe Laskowski und Kato, 1980 und Bode und Huber, 1992). So schützen die in der Leber synthetisierten, humanen Plasma-Proteinase-Inhibitoren α_1 -Antichymotrypsin und α_1 -Proteinase-Inhibitoren das Lungengewebe vor unspezifischem Angriff durch die Proteinasen Kathepsin G bzw. Elastase aus polymorp-kernigen Lymphozyten. Bei einem Ungleichgewicht zwischen Proteinasen und ihren spezifischen Inhibitoren kann es zum Auftreten pathologischer Effekte kommen. Ein übermäßiges Verhältnis von Elastase zu α_1 -Proteinase-Inhibitor erhöht beispielsweise bei Patienten mit genetisch bedingtem Mangel an diesem Faktor das Risiko der Bildung eines Lungenemphysems um ca. 20 bis 30fach gegenüber der Normalbevölkerung (Carrel und Owen, 1980). Bei Rauchern wird die Emphysembildung mittels Oxidation der im reaktiven Zentrum des α_1 -Proteinase-Inhibitors befindlichen Aminosäure Methionin durch im Zigarettenrauch enthaltene Oxidantien begünstigt (Miller und Kuschner, 1969; Ohlsson et al., 1980). Auch im Falle der Infektion mit Gram-negativen Bakterien können deren Endotoxine eine Desintegration von Phagozyten und damit die Ausschüttung lysosomaler Proteinasen verursachen, was durch den erhöhten Verbrauch an Proteinase-Inhibitoren unkontrollierte Gewebsschädigung und Entzündungen verursachen kann. Aus diesem Grund besitzen bestimmte Proteinase-Inhibitoren ein hohes therapeutisches Potential (siehe z.B. Fritz, 1980).

Das der Erfindung zugrunde liegende technische Problem besteht darin, weitere Inhibitoren von Serin-Proteinasen zur Verfügung zu stellen. Des weiteren sollten die für die erfindungsgemäßen Inhibitoren kodierenden Gene bzw. cDNA zur Verfügung gestellt werden.

Spezifisches Merkmal der erfindungsgemäßen Serin-Proteinase-Inhibitoren ist, daß der Serin-Proteinase-Inhibitor eine Domäne mit vier Cysteinen aufweist und sich zwischen dem ersten und einem zweiten Cystein eine Sequenz von 0 bis 20 Aminosäuren

befindet oder der Serin-Proteinase-Inhibitor eine Domäne mit sechs Cysteinen aufweist und sich zwischen dem ersten und zweiten Cystein eine Sequenz von 7 bis 20 Aminosäuren befindet.

Bevorzugterweise befindet sich zwischen einem ersten und einem zweiten Cystein eine Sequenz von 13 Aminosäuren und/oder zwischen einem zweiten und einem dritten Cystein eine Sequenz von 18 Aminosäuren und/oder zwischen einem dritten und vierten Cystein eine Sequenz von 2 Aminosäuren.

Es wird insbesondere bevorzugt, daß die Sequenz zwischen einem ersten und zweiten Cystein ausgewählt wird aus

HEFQAFMKNGKLF,	SEYRKSRKNGRLF,
DDFKKGERDGDFI,	SEFRDQVRNGTLI,
SAFRPFVRNGRLG,	SEYRHYVRNGRLP,
KEYEKQVRNGRLF,	DEFRRLLQNGKLF,
SQYQQNQAKNGILF,	AEYREQMKNGRLS oder
NEYRKLVRNGKLA,	DEFRSQMKNGKLI

und/oder die Sequenz zwischen einem zweiten und dritten Cystein ausgewählt wird aus

PQDKKFFQSLDGIMFINK,	TRENDPIQGPDGKMHGNT,
TRENDPVLPDGKTHGNK,	TREHNPVRGPDGKMHGNK,
TRESDPVRGPDGKMHGNK,	TRENDPIEGLDGKIHGNT,
TRENDPIRGPDGKMHGNL,	TRENDPVGPDGKTHGNK,
TRENDPIQGPDGKVHGNT,	TRESDPVRDADGKSYNQ oder
	TRESDPVRGPDGKTHGNK

und/oder die Sequenz zwischen einem dritten und vierten Cystein ausgewählt wird aus

AT, AL, AM, SM oder TM.

Besonders bevorzugt wird, daß der erfindungsgemäße Serin-Proteinase-Inhibitor einer der Formeln

R₁-C-HEFQAFMKNGKLF-C-PQDKKFFQSLDGIMFINK-C-AT-C-R₂
R₁-C-DDFKKGERDGDFI-C-PDYYEAVCGTDGKTYDNR-C-AL-C-R₂
R₁-C-SAFRPFVRNGRLG-C-TRENDPVLGPDGKTHGNK-C-AM-C-R₂
R₁-C-KEYEKQVRNGRLF-C-TRESDPVRGPDGRMHGNK-C-AL-C-R₂
R₁-C-SQYQNQAKNGILF-C-TRENDPIRGPDGKMHGNL-C-SM-C-R₂
R₁-C-NEYRKLVNGKLA-C-TRENDPIQGPDGKVGNT-C-SM-C-R₂
R₁-C-SEYRKSRSRNGRLF-C-TRENDPIQGPDGKMHGNT-C-SM-C-R₂
R₁-C-SEFRDQVRNGTLI-C-TREHNPVRGPDGKMHGNK-C-AM-C-R₂
R₁-C-SEYRHYVRNGRLP-C-TRENDPIEGLDGKIHGNT-C-SM-C-R₂
R₁-C-DEFRLLQNGKLF-C-TRENDPVRGPDGKTHGNK-C-AM-C-R₂
R₁-C-AEYREQMKNGRLS-C-TRESDPVRDADGKSYNNQ-C-TM-C-R₂
R₁-C-DEFRSQMKNGKLI-C-TRESDPVRGPDGKTHGNK-C-TM-C-R₂,

worin R₁ NH₂, eine Aminosäure oder ein Peptid mit bis zu 100 Aminosäuren ist und R₂ COOH, CONH₂, eine Aminosäure oder ein Peptid mit bis zu 100 Aminosäuren ist, entspricht.

Es ist weiterhin bevorzugt, daß die Serin-Proteinase ein oder mehrere Disulfidbrücken aufweist. Dabei ist besonders bevorzugt, daß er zwischen dem ersten und vierten Cystein und/oder dem zweiten und dritten Cystein eine Disulfidbrücke aufweist oder daß er zwischen dem ersten und fünften Cystein und/oder dem zweiten und vierten Cystein und/oder dem dritten und sechsten Cystein eine Disulfidbrücke aufweist.

Bevorzugte Vertreter der erfindungsgemäßen Serin-Proteinase-Inhibitoren sind die Verbindungen VAKTI-I, VAKTI-II, HF 6479 und HF 7665. Die Aminosäuresequenzen der Proteinase-Inhibitoren sind wiedergegeben in den Aminosäuresequenzen der beigefügten Figuren 1 bis 3.

Aus den Figuren 1 bis 3 lassen sich neben der Aminosäuresequenz der erfindungsgemäß bevorzugten Verbindungen auch weitere Informationen bezüglich der cDNA, die für die erfindungsgemäßen Verbindungen kodiert, entnehmen. Insbesondere werden die entsprechenden Motive und Primer hybridisierenden Stellen angegeben.

M 04.02.99

- 5 -

Die erfindungsgemäße Verbindung VAKTI-I weist eine Masse von 6.479 Dalton auf, diejenige von VAKTI-II beträgt 7.665 Dalton.

Erfindungsgemäß beansprucht wird auch eine cDNA, kodierend für die erfindungsgemäßen Verbindungen, insbesondere eine cDNA mit der Nucleinsäuresequenz gemäß Figuren 1 bis 3.

Die erfindungsgemäßen Verbindungen sind als Arzneimittel geeignet. Gegebenenfalls werden sie zusammen mit pharmazeutisch verträglichen Trägerstoffen appliziert.

Die erfindungsgemäßen Arzneimittel enthaltend die erfindungsgemäßen Proteinase-Inhibitoren werden vorzugsweise in Mengen von 1 bis 100 mg/kg Körpergewicht des Patienten verabreicht. Als Verabreichungsform kommen alle galenischen Zubereitungen für Peptidwirkstoffe in Frage. Die Arzneimittel enthaltend Nucleinsäuren gemäß der Erfindung werden vorzugsweise in Mengen von 0,1 bis 100 mg/kg Körpergewicht eines entsprechenden Patienten verabreicht. Hier kommen als galenische Verabreichungsformen solche in Betracht, die zur Applikation von Nucleinsäuren geeignet sind, ohne daß die Nucleinsäuren vor Erreichen des Wirkortes durch Stoffwechselinflüsse unwirksam gemacht werden. Als galenische Verabreichungsform können z.B. Liposomen eingesetzt werden, in denen die Nucleinsäuren befindlich sind.

Die erfindungsgemäßen Verbindungen kommen insbesondere zur Behandlung von akuten oder chronischen Cervixentzündungen, Entzündungen der Bartholinschen Drüsen oder anderer vaginaler Bereiche, Tonsillitis, Pharyngitis und Laryngitis, mit exzessiver Schleimbildung verbundener akut oder chronisch entzündlicher Prozesse und sich daraus ergebender akuter Notsituationen, postoperativer Blutungen aufgrund Hyperfibrinolyse sowie zur Prophylaxe der Lungenemboliebildung bei α_1 -Proteinase-Inhibitormangel in Betracht.

Die erfindungsgemäßen Verbindungen können bei Mangel an Serin-Proteinase-Inhibitoren verabreicht werden, um endogene Defizite

auszugleichen. Die Nucleinsäuren können, direkt oder an geeignete Vehikel gekoppelt, auch zum Einsatz in der Gentherapie gelangen. Als geeignete Vektoren kommen insbesondere attenuierte Adenoviren, in die entsprechenden Gene inkorporiert werden, in Frage.

Die erfindungsgemäßen Polypeptide, insbesondere VAKTI-I und VAKTI-II, können zur Herstellung von Antikörpern oder Antikörperfragmenten dienen. Diese werden in einfacher Weise durch Immunisierung geeigneter Säuger hergestellt. Durch an sich bekannte Operationen können die Antikörper auch humanisiert werden, so daß diese Antikörper ebenfalls zum therapeutischen Einsatz gelangen können. Antikörper oder Antikörperfragmente können dann zur Regulation von Erkrankungen eingesetzt werden, bei denen die Proteinase-Inhibitoren pathologisch exprimiert werden. Ebenso können zu den erfindungsgemäßen Nucleinsäuren komplementäre Antisense-Nucleinsäuren zum therapeutischen Einsatz bei Überexpression der Proteinase-Inhibitor-Genen eingesetzt werden.

Die erfindungsgemäßen Verbindungen sind in einfacher Weise durch an sich bekannte Methoden der Peptid- bzw. Nucleotidsynthese herstellbar. Einer gentechnischen Herstellung der Verbindungen steht ebenfalls nichts im Wege.

Der Fachmann erkennt, daß bei den Polypeptiden gemäß der Erfindung auch Fragmente verwendet werden können, sofern sie die inhibitorischen Eigenschaften der Serin-Proteinase-Inhibitoren beibehalten. Das Auffinden solcher Fragmente ist dem Fachmann bekannt. So erfolgt dies beispielsweise durch gezielte enzymatische Spaltung der erfindungsgemäßen Verbindungen. Es können auch in den Seitenketten modifizierte Aminosäuren eingesetzt werden. Auch N- und C-terminal modifizierte Polypeptide kommen in Betracht. Insbesondere können phosphorylierte, glycosylierte, methylierte, acetylierte oder in ähnlicher Weise modifizierte Polypeptide eingesetzt werden, sofern sie die Wirkung der Serin-Proteinase-Inhibitoren nicht relevant beeinträchtigen.

Bei den Nucleinsäuren gemäß der Erfindung kommen auch Derivate in Betracht, die je nach Codon Usage modifizierte Tripletstrukturen aufweisen. Desweiteren sind als Nucleinsäuren gemäß der Erfindung auch solche zu verstehen, die durch Nucleasen gegenüber den nativen Verbindungen weniger stark abgebaut werden, beispielsweise die entsprechenden SODN-Derivate, die in der Antisense-Technologie üblicherweise eingesetzt werden, um die Antisense-Strukturen gegenüber enzymatischen Angriffen stabiler auszugestalten.

Auch mit den Polypeptiden homologe Strukturen kommen in Betracht. Dies sind insbesondere Polypeptidstrukturen, bei denen Aminosäuren ausgetauscht sind. So können beispielsweise konservative Aminosäuresubstitutionen in hochkonservierten Regionen wie folgt berücksichtigt werden: Jede Isoleucin-, Valin- und Leucin-Aminosäure kann gegen eine andere dieser Aminosäuren ausgetauscht sein, Aspartat kann gegen Glutamat und umgekehrt ausgetauscht sein, Glutamin gegen Asparagin und umgekehrt, Serin gegen Trionin und umgekehrt. Konservative Aminosäuresubstitutionen in weniger hochkonservierten Regionen können wie folgt sein: Jede der Aminosäuren Isoleucin, Valin und Leucin gegen jede andere Aminosäuren, Aspartat gegen Glutamat und umgekehrt, Glutamin gegen Asparagin und umgekehrt, Serin gegen Theonin und umgekehrt, Glycin gegen Alanin und umgekehrt, Alanin gegen Valin und umgekehrt, Methionin gegen jede der Aminosäuren Leucin, Isoleucin oder Valin, Lysin gegen Arginin und umgekehrt, eine der Aminosäuren Aspartat oder Glutamat gegen eine der Aminosäuren Arginin oder Lysin, Histidin gegen eine der Aminosäuren Arginin oder Lysin, Glutamin gegen Glutamat und umgekehrt und Asparagin gegen Aspartat und umgekehrt.

Die Wirkungsweise der erfindungsgemäßen Peptide wird durch das folgende Beispiel erläutert.

Beispiel

Messung der Proteinase Inhibition durch HF 7665

Meßansatz:

84 μ l Meßpuffer (0,1 M HEPES, pH 7.5; 0,5 M NaCl₂)
1 μ l Trypsin (1 mg/ml in 1 mM HCl, 20 mM CaCl₂)
5 μ l L-BABNA (6 mg/ml α -Benzoyl-L-Arginine-p-Nitroanilide Hydrochloride)
10 μ l Proteaseinhibitor (10 μ M bzw. 75 μ g/ml HF 7665 in H₂O).

Die Reaktion wurde durch Zugabe des chromogenen Substrates gestartet und der Substratumsatz mittels Photometer bei $\lambda = 405$ nm verfolgt. Nach ca. fünf Minuten wurden 10 μ l Proteaseinhibitor bzw. entsprechende Kontrollen dazugegeben und der weitere Extinktionsverlauf beobachtet.

Es konnte gezeigt werden, daß HF 7665 in einer Endkonzentration von ca. 1 μ M bzw. 7,5 μ g/ml einen inhibitorischen Effekt auf Trypsin besitzt. Kontrollversuche mit entsprechenden Mengen an BSA (7,5 μ g/ml) und Acetonitril/TFA (0,8% ACN/0,001% TFA) zeigten keine Trypsininhibierung. Weiterhin konnte kein inhibitorischer Effekt von HF 7665 auf Chymotrypsin bei einem ähnlichen Test beobachtet werden.

Figur 4 zeigt, daß sich nach Zugabe von HF 7665 der Substratumsatz durch Trypsininhibierung um ca. 30% vermindert.

Patentansprüche

1. Serin-Proteinase-Inhibitor, dadurch gekennzeichnet, daß der Serin-Proteinase-Inhibitor eine Domäne mit vier Cysteinen aufweist und sich zwischen einem ersten und einem zweiten Cystein eine Sequenz von 0 bis 20 Aminosäuren befindet oder der Serin-Proteinase-Inhibitor eine Domäne mit sechs Cysteinen aufweist und sich zwischen einem ersten und zweiten Cystein eine Sequenz von 7 bis 20 Aminosäuren befindet.
2. Serin-Proteinase-Inhibitor gemäß Anspruch 1, dadurch gekennzeichnet, daß sich zwischen dem ersten und dem zweiten Cystein der Domäne eine Sequenz von 13 Aminosäuren befindet.
3. Serin-Proteinase-Inhibitor gemäß Anspruch 1 und/oder 2, dadurch gekennzeichnet, daß sich zwischen dem zweiten und dem dritten Cystein der Domäne eine Sequenz von 18 Aminosäuren befindet.
4. Serin-Proteinase-Inhibitor gemäß einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß sich zwischen dem dritten und vierten Cystein der Domäne eine Sequenz von 2 Aminosäuren befindet.
5. Serin-Proteinase-Inhibitor gemäß einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß die Sequenz der Domäne zwischen dem ersten und dem zweiten Cystein ausgewählt wird aus

HEFQAFMKNGKLF, SEYRKSRKNGRLF,
DDFKKGERDGDFI, SEFRDQVRNGTLI,

SAFRPFVRNGRLG, SEYRHYVRNGRLP,
KEYEKQVRNGRLF, DEFRRLLQNGKLF,
SQYQNQAKNGILF, AEYREQMKNGRLS oder
NEYRKLVRNGKLA, DEFRSQMKNGKLI.

6. Serin-Proteinase-Inhibitor gemäß einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die Sequenz zwischen dem zweiten und dem dritten Cystein der Domäne ausgewählt wird aus

PQDKKFFQSLDGIMFINK, TRENDPIQGPDGKMHGNT,
TRENDPVLGPDGKTHGNK, TREHNPVRGPDGKMHGNK,
TRESDPVRGPDGRMHGNK, TRENDPIEGLDGKIHGNT,
TRENDPIRGPDGKMHGNL, TRENDPVRGPDGKTHGNK,
TRENDPIQGPDGKVHGNT, TRESDPVRDADGKSYNNQ oder
TRESDPVRGPDGKTHGNK.

7. Serin-Proteinase-Inhibitor gemäß einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, daß die Sequenz zwischen dem dritten und dem vierten Cystein der Domäne ausgewählt wird aus

AT, AL, AM, SM oder TM.

8. Serin-Proteinase-Inhibitor gemäß einem der Ansprüche 1 bis 7, mit einer der Formeln

R₁-C-HEFQAFMKNGKLF-C-PQDKKFFQSLDGIMFINK-C-AT-C-R₂
R₁-C-DDFKKGERDGDFI-C-PDYYEAVCGTDGKTYDNR-C-AL-C-R₂
R₁-C-SAFRPFVRNGRLG-C-TRENDPVLGPDGKTHGNK-C-AM-C-R₂
R₁-C-KEYEKQVRNGRLF-C-TRESDPVRGPDGRMHGNK-C-AL-C-R₂
R₁-C-SQYQNQAKNGILF-C-TRENDPIRGPDGKMHGNL-C-SM-C-R₂
R₁-C-NEYRKLVRNGKLA-C-TRENDPIQGPDGKVHGNT-C-SM-C-R₂
R₁-C-SEYRKSRSRKNGLF-C-TRENDPIQGPDGKMHGNT-C-SM-C-R₂
R₁-C-SEFRDQVRNGTLI-C-TREHNPVRGPDGKMHGNK-C-AM-C-R₂
R₁-C-SEYRHYVRNGRLP-C-TRENDPIEGLDGKIHGNT-C-SM-C-R₂
R₁-C-DEFRRLLQNGKLF-C-TRENDPVRGPDGKTHGNK-C-AM-C-R₂

13. Verwendung der für die Verbindungen nach mi der Ansprüche 1 bis 9 kodierenden Nucleinsä nucleotids zur Herstellung eines Arzneimitt in der Gentherapie zur Heilung und Prophylakungen gemäß Anspruch 12.
14. Antikörper oder Antikörper-Fragmente geg Verbindungen nach einem der Ansprüche 1 b:
15. Poly- oder Oligonucleotide, die mit Bereich entsprecher RNA unter stringenten Bedin sieren und gegebenenfalls die Expressi Bereiche der für die Verbindungen nach Ar codierenden Gene verhindern (Antisense Ve
16. Diagnostikum enthaltend mindestens eine d gemäß Anspruch 14 oder 15.
17. Arzneimittel enthaltend mindestens eine de chen 14 und/oder 15 genannten Verbindungen wirksamen Mengen.
18. Verwendung der Verbindungen gemäß Anspruc zur Herstellung eines Arzneimittels zur Erkrankungen, die mit einer zu hohen Expressi onen gemäß mindestens einem der Ansprü zu hohen Aktivität der für die Verbindung chen 1 bis 9 codierenden Bereichen verbu
19. DNA kodierend für die in den Ansprüchen 1 Verbindungen und/oder RNA, die in der Tr Translation der in den Ansprüchen 1 bis bindungen beteiligt ist.

Figur 1

VAKTI-2 cDNA und Translation

Aminosäuresequenz

Frame 2

ATG	CAT	GGA	GTG	GAC	CTG	TAG	GCG	ACT	TGC	ATC	GTC	TTC	AAC	ATG	AA
10															46

T	V	S	V	L	L	P	L	A	L	C	L	I	Q	D	A	A	S
ACA	GTG	TCA	GTG	CTT	CTG	CCC	TTG	GCT	CTT	TGC	CTC	ATA	CAA	GAT	GCT	GCC	AG
64		73			82				91				100				10

Repeat 1

E	D	Q	E	M	C	H	E	F	Q	A	F	M	K	N	G	K	L
GAA	GAT	CAG	GAA	ATG	TGC	CAT	GAA	TTT	CAG	GCA	TTT	ATG	AAA	AAT	GGA	AAA	CT
124		133			142				151			160					16

P	Q	D	K	K	F	F	Q	S	L	D	G	I	M	F	I	N	K
CCC	CAG	GAT	AAG	AAA	TTT	TTT	CAA	AGT	CTT	GAT	GGA	ATA	ATG	TTC	ATC	AAT	AA
184		193			202				211			220					22

T	C	K	M	I	L	E	K	E	A	K	S	Q	K	R	A	R	H
ACG	TGC	AAA	ATG	ATA	CTG	GAA	AAA	GAA	GCA	AAA	TCA	CAG	AAG	AGG	GCC	AGG	CA
244		253			262				271			280					28

Typische Kazal-Domäne

R	A	P	K	A	T	A	P	T	E	L	N	C	D	D	F	K	K
AGA	GCT	CCC	AAG	GCT	ACT	GCC	CCA	ACA	GAG	CTG	AAT	TGT	GAT	GAT	TTT	AAA	AA
304		313			322				331			340					34

R	D	G	D	F	I	C	P	D	Y	Y	E	A	V	C	G	T	D
AGA	GAT	GGG	GAT	TTT	ATC	TGT	CCT	GAT	TAT	TAT	GAA	GCT	GTT	TGT	GGC	ACA	GA
364		373			382				391			400					40

T	Y	D	N	R	C	A	L	C	A	E	N	A	K	T	G	S	Q
ACA	TAT	GAC	AAC	AGA	TGT	GCA	CTG	TGT	GCT	GAG	AAT	GCG	AAA	ACC	GGG	TCC	CA
424		433			442				451			460					46

Repeat

V	K	S	E	G	E	C	K	S	S	N	P	E	Q	D	V	C	S
GTA	AAA	AGT	GAA	GGG	GAA	TGT	AAG	AGC	AGT	AAT	CCA	GAG	QAG	GAT	GTA	TGC	AG
484		493			502				511			520					52

R	P	F	V	R	N	G	R	L	G	C	T	R	E	N	D	P	V
CGG	CCC	TTT	GTT	AGA	AAT	GGA	AGA	CTT	GGA	TGC	ACA	AGG	GAA	AAT	GAT	CCT	GT
544		553			562				571			580					58

P	D	G	K	T	H	G	N	K	C	A	M	C	A	E	L	F	L
CCT	GAT	GGG	AAG	ACG	CAT	GGC	AAT	AAG	TGT	GCA	ATG	TGT	GCT	GAG	CTG	TTT	TT
604		613			622				631			640					64

A	E	N	A	K	R	E	G	E	T	R	I	R	R	N	A	E	K
GCT	GAA	AAT	GCC	AAG	CGA	GAG	GGT	GAA	ACT	AGA	ATT	CGA	CGA	AAT	GCT	GAA	AA
664		673			682				691			700					70

Repeat 3

C	K	E	Y	E	K	Q	V	R	N	G	R	L	F	C	T	R	E
TGC	AAG	GAA	TAT	GAA	AAA	CAA	GTG	AGA	AAT	GGA	AGG	CTT	TTT	TGT	ACA	CGG	GAC
724		733			742				751			760					76

P	V	R	G	P	D	G	R	M	H	G	N	K	C	A	L	C	A
CCA	GTC	CGT	GGC	CCT	GAC	GGC	AGG	ATG	CAT	GGC	AAC	AAA	TGT	GCC	CTG	TGT	GCT
784		793			802				811			820					82

F	K	R	R	F	S	E	E	N	S	K	T	D	Q	N	L	G	K
TTC	AAG	CGG	CGT	TTT	TCA	GAG	GAA	AAC	AGT	AAA	ACA	GAT	CAA	AAT	TTG	GGA	AA
844		853			862				871			880					88

Repeat 4

E	K	T	K	V	K	R	E	I	V	K	L	C	S	Q	Y	Q	N
GAA	AAA	ACT	AAA	GTT	AAA	AGA	GAA	ATT	GTG	AAA	CTC	TGC	AGT	CAA	TAT	CAA	AA
904		913			922				931			940					94

K	N	G	I	L	F	C	T	R	E	N	D	P	I	R	G	P	D
AAG	AAT	GGA	ATA	CTT	TTC	TGT	ACC	AGA	GAA	AAT	GAC	CCT	ATT	CGT	GGT	CCA	GAT
964		973			982				991			1000					100

M	H	G	N	L	C	S	M	C	Q	V	Y	F	Q	N	E		
ATG	CAT	GGC	AAC	TTG	TGT	TCC	ATG	TGT	CAA	GTC	TAC	TTC	CAA	GCA	GAA	AAT	
1024		1033			1042				1051			1060					106

104-03

Repeat 10

D	E	C	A	E	Y	R	E	Q	M	K	N	G	R	L	S	C	T	R	E
GAC	GAA	TGT	GCT	GAG	TAT	CGG	GAA	CAA	ATG	AAA	AAT	GGA	AGA	CTC	AGC	TGT	ACT	CGG	G
2164			2173			2182			2191			2200					2209		

S	D	P	V	R	D	A	D	G	K	S	Y	N	N	Q	C	T	M	C	*
AGT	GAT	CCT	GTA	CGT	GAT	GCT	GAT	GGC	AAA	TCG	TAC	AAC	AAT	CAG	TGT	ACC	ATG	TGT	A
2224			2233			2242			2251			2260					2269		

A	K	L	E	R	E	A	E	R	K	N	E	Y	S	R	S	R	S	N	G
GCA	AAA	TTG	GAA	AGA	GAA	GCA	GAG	AGA	AAA	AAT	GAG	TAT	TCT	CGC	TCC	AGA	TCA	AAT	G
2284			2293			2302			2311			2320					2329		

T	G	S	E	S	G	K	D	T	C	D	E	F	R	S	Q	M	K	N	G
ACT	GGA	TCA	GAA	TCA	GGG	AAG	GAT	ACA	TGT	GAT	GAG	TTT	AGA	AGC	CAA	ATG	AAA	AAT	G
2344			2353			2362			2371			2380					2389		

K	L	I	C	T	R	E	S	D	P	V	R	G	P	D	G	K	T	H	G
AAA	CTT	ATC	TGC	ACT	CGA	GAA	AGT	GAC	CCT	GTC	CGG	GGT	CCA	GAT	GGC	AAG	ACA	CAT	G
2404			2413			2422			2431			2440					2449		

N	K	C	T	M	C	K	E	K	L	E	R	E	A	A	E	K	K	R	K
AAT	AAG	TGT	ACT	ATG	TGT	AAG	GAA	AAA	CTG	GAA	AGG	GAA	GCA	GCT	GAA	AAA	AAA	AGA	A
2464			2473			2482			2491			2500					2509		

R	M	K	T	G	A	I	Q	E	K	G	A	I	Q	E	K	G	A	M	T
AGG	ATG	AAG	ACA	GGA	GCA	ATA	CAG	GAG	AAA	GGA	GCA	ATA	CAG	GAG	AAA	GGA	GCA	ATG	A
2524			2533			2542			2551			2560					2569		

K	R	I	C	V	V	N	F	E	A	C	R	E	M	E	S	L	S	A	P
AAG	AGG	ATC	TGT	GTC	GTG	AAT	TTC	GAA	GCA	TGC	AGA	GAA	ATG	GAA	AGC	TTA	TCT	GCA	C
2584			2593			2602			2611			2620					2629		

E	K	I	T	L	F	E	A	H	M	A	R	C	T	S	I	N	V	L	C
GAG	AAA	ATA	ACC	CTG	TTC	GAG	GCC	CAT	ATG	GCA	AGA	TGC	ACA	TCA	ATA	ATA	AT	GTG	CTA
2644			2653			2662			2671			2680					2689		

V	R	A	S	L	I	E	K	L	M	K	E	K	R	K	M	K	R	N	Q
GTC	AGA	GCA	TCT	TTG	ATC	GAG	AAG	CTA	ATG	AAA	GAA	AAA	AGA	AAG	ATG	AAG	AGA	AAT	C
2704			2713			2722			2731			2740					2749		

V	A	S	P	Q	I	M	Q	R	M	S	A	V	N	F	E	T	I	STOP	
GTA	GCA	AGC	CCT	CAA	ATA	ATG	CAA	AGG	ATG	AGT	GCA	GTG	AAT	TTC	GAA	ACT	ATA	TAA	
2764			2773			2782			2791			2800					2809		

ACA	ATG	AAC	TCA	TCT	GCC	CTA	GAG	AGA	ATG	ACC	CAG	TGC	ACG	GTG	CTG	ATG	GAA	AGT	TG
2824			2833			2842			2851			2860					2869		

ATA	CAA	ACA	AGT	GCT	ACA	TGT	GCA	GAG	CTG	TCT	TTC	TAA	CAG	AAG	CTT	TGG	AAA	GGG	C
2884			2893			2902			2911			2920					2929		

AGC	TTC	AAG	AAA	AAC	CAT	CCC	ATG	TTA	GAG	CTT	CTC	AAG	AGG	AAG	ACA	GCC	CAG	ACT	C
2944			2953			2962			2971			2980					2989		

TCA	GTT	CTC	TGG	ATT	CTG	AGA	TGT	GCA	AAG	ACT	ACC	GAG	TAT	TGC	CCA	GGA	TAG	GCT	A
3004			3013			3022			3031			3040					3049		

TTT	GTC	CAA	AGG	ATT	TAA	AGC	CTG	TCT	GTG	GTG	ACG	ATG	GCC	AAA	CCT	ACA	ACA	ATC	C
3064			3073			3082			3091			3100					3109		

GCA	TGC	TCT	GTC	ATG	AAA	ACC	TGA	TAC	GCC	AAA	CAA	ATA	CAC	ACA	TCC	GCA	GTA	CAG	GC
3124			3133			3142			3151			3160					3169		

AGT	GTG	AGG	AGA	GCA	GCA	CCC	CAG	GAA	CCA	CCG	CAG	CCA	GCA	TGC	CCC	CGT	TTG	ACG	AP
3184			3193			3202			3211			3220					3229		

Figur 2

Vollständige VAKTI-2 cDNA-Sequenz

TATGCATGGA GTGGACCTGT AGGCGACTTG CATCGTCTTC AACATGAAGA TAGCCACAGT 61
 GTCAGTGCTT CTGCCCTTGG CTCTTTGCCT CATAACAAGAT GCTGCCAGTA AGAATGAAGA 12
 TCAGGAAATG TGCCATGAAT TTCAGGCATT TATGAAAAAT GGAAAACGT TCTGTCCCCA 18
 GGATAAGAAA TTTTTTCAAA GTCTTGATGG AATAATGTTA ATCAATAAAAT GTGCCACGTG 24
 CAAAATGATA CTGGAAAAAG AAGCAAAATC ACAGAAGAGG GCCAGGCATT TAGCAAGAGC 30
 TCCAAGGCT ACTGCCCAA CAGAGCTGAA TTGTGATGAT TTTAAAAAAG GAGAAAGAGA 3
 TGGGGATTTT ATCTGTCCCTG ATTATTATGA AGCTGTTGTG GGCACAGATG GGAAAACATA 4
 TGACAACAGA TGTGCACTGT GTGCTGAGAA TGCGAAAACC GGGTCCAAA TTGGTGTAAA 4
 AAGTGAAGGG GAATGTAAGA GCAGTAATCC AGAGCAGGAT GTATGCAGTG CTTTCGGCC 1
 CTTTGTCTAGA AATGGAAGAC TTGGATGCAC AAGGGAAAAT GATCCTGTTG TTGGTCCTGA
 TGGGAAGACG CATGGCAATA AGTGTGCAAT GTGTGCTGAG CTGTTTTAA AAGAAGCTGA
 AAATGCCAAG CGAGAGGGTG AACTAGAAT TCGACGAAAT GCTGAAAAGG ATTTTGCAA
 GGAATATGAA AAACAAGTGA GAAATGGAAG GCTTTTTGT ACACGGGAGA GTGATCCAGT
 CCGTGGCCCT GACGGCAGGA TGCATGGCAA CAAATGTGCC CTGTGCTG AAATTTCAA
 GCGCGTTTT TCAGAGGAAA ACAGTAAAAC AGATCAAAAT TTGGGAAAAG CTGAAGAAAA
 AACTAAAGTT AAAAGAGAAA TTGTGAAACT CTGCAGTCAT TATCAAAATC AGGCAAAGAA
 TGGAATACTT TTCTGTACCA GAGAAAATGA CCCTATTGCGT GGTCCAGATG GGAAAATGCA
 TGGCAACTTG TGTTCCATGT GTCAAGTCTA CTTCCAAGCA GAAAATGAAG AAAAGAAAAA
 GGCTGAAGCA CGAGCTAGAA ACAAAAGAGA ATCTGGAAAA GCAACCTCAT ATGCAGAGCT
 TTGCAATGAA TATCGAAAGC TTGTGAGGAA CGGAAAACCTT GCTTGACCA GAGAGAACGA
 CCCTATTGAG GGCCCAGATG GGAAAGTGCA CGGCAACACC TGCTCCATGT GTGAGGTTTT
 TTTCCAAGCA GAAGAAGAAG AAAAGAAAAA GAAGGAAGGC GAATCAAGAA ACAAAAGACA
 ATCTAAGAGT ACAGCTTCCT TTGAGGAGTT GTGTAGTGAA TACCGCAAAT CCAGGAAAAA
 CGGACGGCTT TTTTGACCA GAGAGAATGA CCCCACCCAG GGCCCAGATG GGAAAATGCA
 TGGCAACACC TGCTCCATGT GTGAGGCTT CTTCAACAA GAAGAAAGAG CAAGAGCAA
 GGCTAAAAGA GAAGCTGCAA AGGAAATCTG CAGTGAATTG CGGGACCAAG TGAGGAATGG
 AACACTTATA TGCACCAAGGG AGCATAATCC TGTCCGTGGA CCAGATGGCA AAATGCATGG
 AAACAAGTGT GCCATGTGTG CCAGTGTGTT CAAACTTGAA GAAGAAGAGA AGAAAATG/
 TAAAGAAGAA AAAGGGAAAAG TTGAGGCTGA AAAAGTTAAG AGAGAAGCAG TTCAGGAGC

Figur 3

Vollständige VAKTI-2 Aminosäuresequenz

MKIATVSVLL PLALCLIQDA ASKNEDQEMC HEFQAFMKNG KLFCPQDKKF FQSLDG
NKCATCKMIL EKEAKSQKRA RHLARAPKAT APTELNCDDF KKGERDGDFI CPDYYEA
TDGKTYDNRC ALCAENAKTG SQIGVKSEGE CKSSNPEQDV CSAFRPFVRN GRLGCTF
PVLGPDGKTH GNKCAMCAEL FLKEAENAKR EGETRIRRNA EKDFCKEYEK QVRNGRI
RESDPVRGPD GRMHGNKCAL CAEIFKRRFS EENSKTDQNL GKAEEKTKVK REIVKL
QNQAKNGILF CTRENDPIRG PDGKMHGNLC SMCQVYFQAE NEEKKKAEAR ARNKRE
TSYAECLNEY RKLVRNGKLA CTRENDPIQG PDGKVHGNTC SMCEVFFQAE EEEKKK
SRNKRSKST ASFEELCSEY RKSRSKNGRLF CTRENDPIQG PDGKMHGNTC SMCEAF
ERARAKAKRE AAKEICSEFR DQVRNGTLIC TREHNPVRGP DGKMHGNKCA MCASVF
EEKKNDKEEK GKVEAEKVKR EAVQELCSEY RHYVRNGRLP CTRENDPIEG LDGKII
SMCEAFFQQE AKEKERAEPK AKVKREAEKE TCDEFRRLLQ NGKLFCTREN DPVRG
HGNKCAMCKA VFQKENEERK RKEEEDQRNA AGHGSSGGGG GNTQDECAEY REQMK
CTRESDPVRD ADGKSYNQNC TMCKAKLERE AERKNEYRSRS RSNGTGSESG KDTCD
MKGKPLICTR ESDPVRGPDG KTHGNKCTMC KEKLEREAAE KKRKRMKTGA IQEKG
GAMTKRICVV NFEACREMES LSAPEKITLF EAHMARCTSI NVLCVRASLI EKLMP
KRNQVASPQI MQRMSAVNFE TI

GAC AGG AAG ATT GTT GAA AGC CAT GAG GGA AAA AAT AAA CCC CAG TTT TGA ATC ACC TAC
3244 3253 3262 3271 3280 3289

CTT CAC CAT CTG TAT ATA CAA AGA ATT TTT CGG AGC TTG TTT TAT TTG CTA TAG AAA ACA
3304 3313 3322 3331 3340 3349

ATA CAG AGC TTT TGG GAA TGG AAT CAC TGA TTT TCA GTC TTT TCC ATT TCT TTC CTC CTA
3364 3373 3382 3391 3400 3409

GAA TCT GTG ATC TGA GGG TAT AAA GAC ATT TCC ACC AAG TTT GAG CCC TCA AAA TGT CCT
3424 3433 3442 3451 3460 3469

GAT TAC AAT GCT GTC TGT CCA ACT GCC TGT ^{Polyadenylierungssignal}
TCA ATA AAA GTA AAC TCA GCA GAA AAA....
3484 3493 3502 3511 3520 3529

.....Poly(A)-Tail

Die Positionen der Hämofiltrat-Peptide HF 6479 und HF 7665 sind markiert. Die Repeats und die typische Kazal-Domäne sind überstrichen. Die erwarteten (und für HF 7665 gezeigte) Cystein-Verbrückungen sind durch Markierung mit jeweils gleichen Zeichen gezeigt. Das konservierte Tyrosin (Y) innerhalb der Kazal-Domäne ist durch ein Ausrufungszeichen markiert. Die Größe der gesamten cDNA ohne Poly(A)-Tail beträgt 3527 bp.

Figur 2

Vollständige VAKTI-2 cDNA-Sequenz

TATGCATGGA	GTGGACCTGT	AGGCGACTTG	CATCGTCTTC	AACATGAAGA	TAGCCACAGT	60
GTCAGTGCTT	CTGCCCTTGG	CTCTTTGCCT	CATACAAGAT	GCTGCCAGTA	AGAATGAAGA	120
TCAGGAAATG	TGCCATGAAT	TTCAGGCATT	TATGAAAAAT	GGAAAACGT	TCTGTCCCCA	180
GGATAAGAAA	TTTTTTCAAA	GTCTTGATGG	AATAATGTT	ATCAATAAAT	GTGCCACGTG	240
CAAAATGATA	CTGGAAAAAG	AAGCAAAATC	ACAGAAGAGG	GCCAGGCATT	TAGCAAGAGC	300
TCCAAGGCT	ACTGCCCAA	CAGAGCTGAA	TTGTGATGAT	TTTAAAAAAG	GAGAAAGAGA	360
TGGGGATT	ATCTGTCTG	ATTATTATGA	AGCTGTTGT	GGCACAGATG	GGAAAACATA	420
TGACAACAGA	TGTGCACTGT	GTGCTGAGAA	TGCGAAAACC	GGGTCCAAA	TTGGTGTAAA	480
AAGTGAAGGG	GAATGTAAGA	GCAGTAATCC	AGAGCAGGAT	GTATGCAGTG	CTTTTCGGCC	540
CTTTGTTAGA	AATGGAAGAC	TTGGATGCAC	AAGGGAAAAT	GATCCTGTT	TTGGTCTGTA	600
TGGGAAGACG	CATGGCAATA	AGTGTGCAAT	GTGTGCTGAG	CTGTTTTAA	AAGAAGCTGA	660
AAATGCCAAG	CGAGAGGGTG	AAACTAGAAT	TCGACGAAAT	GCTGAAAAGG	ATTTTGCAA	720
GGAATATGAA	AAACAAGTGA	GAAATGGAAG	GCTTTTTGT	ACACGGAGA	GTGATCCAGT	780
CCGTGGCCCT	GACGGCAGGA	TGCATGGCAA	CAAATGTGCC	CTGTGTGCTG	AAATTTCAA	840
GCGCGT	TCAGAGGAAA	ACAGTAAAC	AGATCAAAT	TTGGAAAAG	CTGAAGAAAA	900
AACTAAAGTT	AAAAGAGAAA	TTGTGAAACT	CTGCAGTC	TATCAAAATC	AGGCAAAGAA	960
TGGAATACTT	TTCTGTACCA	GAGAAAATGA	CCCTATTG	GGTCCAGATG	GGAAAATGCA	1020
TGGCAACTTG	TGTTCCATGT	GTCAAGTCTA	CTTCCAAGCA	GAAAATGAAG	AAAAGAAAAA	1080
GGCTGAAGCA	CGAGCTAGAA	ACAAAAGAGA	ATCTGGAAAA	GCAACCTCAT	ATGCAGAGCT	1140
TTGCAATGAA	TATCGAAAGC	TTGTGAGGAA	CGGAAAAC	GCTTGCACCA	GAGAGAACGA	1200
TCCTATTCA	GGCCCAGATG	GGAAAGTGCA	CGGCAACACC	TGCTCCATGT	GTGAGGTTT	1260
TTTCCAAGCA	GAAGAAGAAG	AAAAGAAAAA	GAAGGAAGGC	GAATCAAGAA	ACAAAAGACA	1320
ATCTAAGAGT	ACAGCTTCCT	TTGAGGAGTT	GTGTAGTGAA	TACCGCAAAT	CCAGGAAAAA	1380
CGGACGGCTT	TTTGACCA	GAGAGAATGA	CCCCATCCAG	GGCCCAGATG	GGAAAATGCA	1440
TGGCAACACC	TGCTCCATGT	GTGAGGCTT	CTTCAACAA	GAAGAAAGAG	CAAGAGCAA	1500
GGCTAAAAGA	GAAGCTGCAA	AGGAAATCTG	CAGTGAATT	CGGGACCAAG	TGAGGAATGG	1560
AACACTTATA	TGCACCAGGG	AGCATAATCC	TGTCCGTGGA	CCAGATGGCA	AAATGCATGG	1620
AAACAAGTGT	GCCATGTGTG	CCAGTGTGTT	CAAACTTGAA	GAAGAAGAGA	AGAAAAATGA	1680
TAAAGAAGAA	AAAGGGAAAG	TTGAGGCTGA	AAAAGTTAAG	AGAGAAGCAG	TTCAGGAGCT	1740

GTGCAGTGAA TATCGTCATT ATGTGAGGAA TGGACGACTC CCCTGTACCA GAGAGAATGA	1800
TCCTATTGAG GGTCTAGATG GGAAAATCCA CGGCAACACC TGCTCCATGT GTGAAGCCTT	1860
CTTCCAGCAA GAAGCAAAAG AAAAAGAAAG AGCTGAACCC AGAGCAAAAG TCAAAAGAGA	1920
AGCTGAAAAG GAGACATGCG ATGAATTCG GAGACTTTG CAAAATGGAA AACTTTCTG	1980
CACAAGAGAA AATGATCCTG TGCCTGGCCC AGATGGCAAG ACCCATGGCA ACAAGTGTGC	2040
CATGTGTAAG GCAGTCTTCC AGAAAGAAAA TGAGGAAAGA AAGAGGAAAG AAGAGGAAGA	2100
TCAGAGAAAT GCTGCAGGAC ATGGTCCAG TGGTGGTGGA GGAGGAAACA CTCAGGACGA	2160
ATGTGCTGAG TATCGGGAAC AAATGAAAAA TGGAAGACTC AGCTGTACTC GGGAGAGTGA	2220
TCCTGTACGT GATGCTGATG GCAAATCGTA CAACAATCAG TGTACCATGT GTAAAGCAAA	2280
ATTGGAAAGA GAAGCAGAGA GAAAAAATGA GTATTCTCGC TCCAGATCAA ATGGGACTGG	2340
ATCAGAATCA GGGAGGATA CATGTGATGA CTTTAGAAGC CAAATGAAAA ATGGAAAACT	2400
TATCTGCACT CGAGAAAGTG ACCCTGTCCG GGGTCCAGAT GGCAAGACAC ATGGTAATAA	2460
GTGTACTATG TGTAAGGAAA AACTGGAAAG GGAAGCAGCT GAAAAAAA GAAAGAGGAT	2520
GAAGACAGGA GCAATACAGG AGAAAGGAGC AATACAGGAG AAAGGAGCAA TGACAAAGAG	2580
GATCTGTGTC GTGAATTCG AAGCATGCAG AGAAATGGAA AGCTTATCTG CACCAGAGAA	2640
AATAACCTG TTCGAGGCC ATATGGCAAG ATGCACATCA ATAAATGTGC TATGTGTCAG	2700
AGCATCTTG ATCGAGAACG TAATGAAAGA AAAAAGAAAG ATGAAGAGAA ATCAAGTAGC	2760
AAGCCCTCAA ATAATGCAA GGATGAGTGC AGTGAATTTC GAAACTATAT AAGGAACAAT	2820
GAACTCATCT GCCCTAGAGA GAATGACCCA GTGCACGGTG CTGATGGAAA GTTCTATACA	2880
AACAAGTGCT ACATGTGCAG AGCTGTCTTT CTAACAGAAG CTTGGAAAG GGCAAAGCTT	2940
CAAGAAAAAC CATCCCATGT TAGAGCTTCT CAAGAGGAAG ACAGCCCAGA CTCTTTCAGT	3000
TCTCTGGATT CTGAGATGTG CAAAGACTAC CGAGTATTGC CCAGGATAGG CTATCTTGT	3060
CCAAAGGATT TAAAGCCTGT CTGTGGTGAC GATGGCCAAA CCTACAACAA TCCTTGCATG	3120
CTCTGTGATG AAAACCTGAT ACGCCAAACA AATACACACA TCCGCAGTAC AGGGAAGTGT	3180
GAGGAGAGCA GCACCCCCAGG AACCAACCGCA GCCAGCATGC CCCCCTTGA CGAATGACAG	3240
GAAGATTGTT GAAAGCCATG AGGGAAAAAA TAAACCCAG TTTTGAATCA CCTACCTTCA	3300
CCATCTGTAT ATACAAAGAA TTTTCGGAG CTTGTTTAT TTGCTATAGA AAACAATACA	3360
GAGCTTTGG GAATGGAATC ACTGATTTTC AGTCTTTCC ATTTCTTCC TCCTAGAATC	3420
TGTGATCTGA GGGTATAAAG ACATTTCCAC CAAGTTGAG CCCTCAAAAT GTCCTGATTA	3480
CAATGCTGTC TGTCCAATG CCTGTTCAAT AAAAGTAAAC TCAGCAGAAA AA	3532

Figur 3

Vollständige VAKTI-2 Aminosäuresequenz

MKIATVSVLL	PLALCLIQDA	ASKNEDQEMC	HEFQAFMKNG	KLFCPQDKKF	FQSLDGIMFI	60
NKCATCKMIL	EKEAKSQKRA	RHLARAPKAT	APTELNCDDF	KKGERDGDFI	CPDYYEAVCG	120
TDGKTYDNRC	ALCAENAKTG	SQIGVKSEGE	CKSSNPEQDV	CSAFRPFVRN	GRLGCTREND	180
PVLGPDGKTH	GNKCAMCAEL	FLKEAENAKR	EGETRIRRNA	EKDFCKEYEK	QVRNGLFCT	240
RESDPVRGPD	GRMHGNKCAL	CAEIFKRRFS	EENSKTDQNL	GKAEEKTKVK	REIVKLCSQY	300
QNQAKNGILF	CTRENDPIRG	PDGKMHGNLC	SMCQVYFQAE	NEEKKKAEAR	ARNKRESGKA	360
TSYAECLNEY	RKLVRNGKLA	CTRENDPIQG	PDGKVHGNTC	SMCEVFFQAE	EEEKKKKEGE	420
SRNKRQSKST	ASFEELCSEY	RKSRKNGRLF	CTRENDPIQG	PDGKMHGNTC	SMCEAFFQQE	480
ERARAKAKRE	AAKEICSEFR	DQVRNGTLIC	TREHNPVRGP	DGKMHGNKCA	MCASVFKLEE	540
EEKKNDKEEK	GKVEAEKVKR	EAVQELCSEY	RHYVRNGRLP	CTRENDPIEG	LDGKIHGNTC	600
SMCEAFFQQE	AKEKERAEPK	AKVKREAKE	TCDEFRLLLQ	NGKLFCTREN	DPVRGPDGKT	660
HGNKCAMCKA	VFQKENEERK	RKEEEDQRNA	AGHGSSGGGG	GNTQDECAEY	REQMKNGRLS	720
CTRESDPVRD	ADGKSYNQNC	TMCKAKLERE	AERKNEYRSRS	RSNGTGSESG	KDTCDFRSQ	780
MKNGKLICTR	ESDPVRGPDG	KTHGNKCTMC	KEKLEREAAE	KKRKRMKTGA	IQEKGAIQEK	840
GAMTKRICVV	NFEACREMES	LSAPEKITLF	EAHMARCTSI	NVLCVRASLI	EKLMKEKRKM	900
KRNQVASPQI	MQRMSAVNFE	TI				

Figur 4

